



**University of  
Zurich<sup>UZH</sup>**

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2017

---

## **Hat Computational Psychiatry Relevanz für die klinische Praxis der Psychiatrie?**

Stephan, K E ; Siemerikus, J ; Bischof, M ; Haker, H

**Abstract:** Computational Psychiatry (CP) ist ein junges Wissenschaftsfeld an der Schnittstelle zwischen der klinischen Psychiatrie und den mathematischen Neurowissenschaften, das sich in den letzten Jahren zu entfalten begonnen hat. Dieser Artikel widmet sich den möglichen klinischen Implikationen dieser jungen Disziplin. Wir (i) beginnen mit einer kurzen Übersicht über die Geschichte, Ziele und Inhalte der CP, (ii) beschreiben die zentralen Themen, Modelle und Theorien der CP, (iii) untersuchen die Relevanz und das Potenzial modell-basierter diagnostischer Tests (computational assays) für die Lösung zentraler Probleme in der klinischen Psychiatrie, und (iv) stellen zukünftige Herausforderungen und Chancen der CP dar. Computational Psychiatry (CP) is a young field of research at the interface between clinical psychiatry and mathematical neurosciences, which has emerged during recent years. This article describes the possible clinical implications of this young discipline. We (i) begin with a short review on the history, aims, and contents of CP, (ii) delineate the relevance, central themes, models, and theories of CP, (iii) investigate the relevance and the potential of model-based diagnostic tests (computational assays) for the solution of central problems of clinical psychiatry, and (iv) point out future challenges and opportunities of CP.

DOI: <https://doi.org/10.1024/1661-4747/a000296>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-133499>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Stephan, K E; Siemerikus, J; Bischof, M; Haker, H (2017). Hat Computational Psychiatry Relevanz für die klinische Praxis der Psychiatrie? *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 65(1):9-19.  
DOI: <https://doi.org/10.1024/1661-4747/a000296>

# Hat Computational Psychiatry Relevanz für die klinische Praxis der Psychiatrie?

Klaas Enno Stephan<sup>1</sup>, Jakob Siemerikus<sup>1,2</sup>, Martin Bischof<sup>1,2</sup> und Helene Haker<sup>1</sup>

<sup>1</sup> Translational Neuromodeling Unit (TNU), Institut für Biomedizinische Technik, Universität Zürich & ETH Zürich, Schweiz

<sup>2</sup> Klinik für Psychiatrie, Psychotherapie und Psychosomatik, Psychiatrische Universitätsklinik Zürich, Schweiz

**Zusammenfassung:** Computational Psychiatry (CP) ist ein junges Wissenschaftsfeld an der Schnittstelle zwischen der klinischen Psychiatrie und den mathematischen Neurowissenschaften, das sich in den letzten Jahren zu entfalten begonnen hat. Dieser Artikel widmet sich den möglichen klinischen Implikationen dieser jungen Disziplin. Wir (i) beginnen mit einer kurzen Übersicht über die Geschichte, Ziele und Inhalte der CP, (ii) beschreiben die zentralen Themen, Modelle und Theorien der CP, (iii) untersuchen die Relevanz und das Potenzial modell-basierter diagnostischer Tests (computational assays) für die Lösung zentraler Probleme in der klinischen Psychiatrie, und (iv) stellen zukünftige Herausforderungen und Chancen der CP dar.

**Schlüsselwörter:** komputationale Psychiatrie, komputationale Neurowissenschaft, Translational Neuromodeling, komputationale diagnostische Tests

## Computational Psychiatry: is it relevant for psychiatric practice?

**Abstract:** Computational Psychiatry (CP) is a young field of research at the interface between clinical psychiatry and mathematical neurosciences, which has emerged during recent years. This article describes the possible clinical implications of this young discipline. We (i) begin with a short review on the history, aims, and contents of CP, (ii) delineate the relevance, central themes, models, and theories of CP, (iii) investigate the relevance and the potential of model-based diagnostic tests (computational assays) for the solution of central problems of clinical psychiatry, and (iv) point out future challenges and opportunities of CP.

**Keywords:** computational psychiatry, computational neuroscience, translational neuromodeling, computational assays

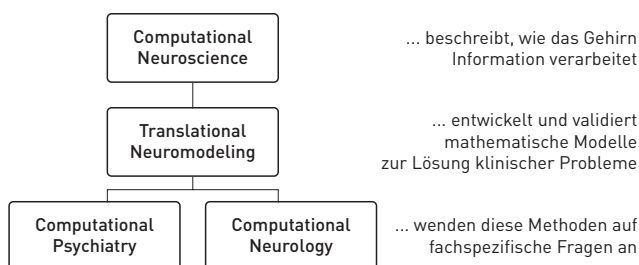
## Was ist Computational Psychiatry?

Die Computational Psychiatry (CP) ist eine junge Disziplin an der Schnittstelle zwischen der klinischen Psychiatrie und den mathematischen Neurowissenschaften.

Je nach Auslegung des Begriffs „computational“ (deutsch: „komputational“), kann CP breit oder eng definiert werden, wie sich in einer Reihe jüngst erschienener Übersichtsartikel zeigt (Corlett & Fletcher, 2014; Friston, Stephan, Montague & Dolan, 2014; Huys, Maia & Frank, 2016; Montague, Friston & Dayan, 2012; Stephan & Mathys, 2014; Wang & Krystal, 2014). In diesem Artikel verwenden wir eine eher breite Definition und verstehen unter CP die Anwendung mathematischer und algorithmischer Methoden, um Fragen der Nosologie, Diagnostik, Prognose und Therapieprädiktion in der Psychiatrie zu untersuchen. Allerdings konzentrieren wir uns hier auf die

Entwicklung und Anwendung mathematischer Modelle, die ein mechanistisches Verständnis von Prozessen und Daten anstreben – insbesondere sogenannte „generative Modelle“ – um klinische Probleme in der Psychiatrie zu lösen. Diese Definition grenzt die CP von deskriptiven Ansätzen ab, wie zum Beispiel statistische Mustererkennung und maschinelles Lernen (Machine Learning).

Historisch ist die CP aus der Computational Neuroscience entstanden, die sich mit der mathematischen Untersuchung von Informationsverarbeitung auf allen Ebenen des Nervensystems (z.B. einzelne Neurone, Netzwerke, aber auch Kognition) beschäftigt, und stellt einen speziellen Anwendungsfall des Translational Neuromodeling (TN) dar. TN entwickelt mathematische Modelle für klinische Problemlösungen in fachübergreifender Weise und schliesst neben der CP auch die Schwesterdisziplin, Computational Neurology, ein (Abbildung 1).



**Abbildung 1.** Eine Taxonomie komputationaler neurowissenschaftlicher Disziplinen und ihr klinischer Bezug.

beobachtbaren Krankheitsbilder entstehen. Entsprechend sind die pathophysiologischen Konzepte hinter einer Psychopharmakotherapie zumeist vage und die Medikationsselektion mehr von klinischer Intuition denn von einer zuverlässigen Prädiktion des individuellen Ansprechens geleitet.

Sowohl für die konzeptionelle Formulierung von pathophysiologischen Mechanismen hinter den heutigen heterogenen Krankheitsentitäten, als auch für die Entwicklung von Methoden zur Detektion individueller Krankheitsprozesse bietet die CP neue Ansätze.

## Wozu Computational Psychiatry?

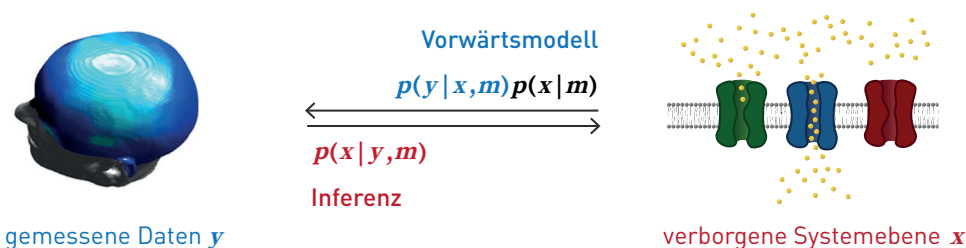
Viele ungelöste klinische Probleme in der Psychiatrie hängen mit der heutigen Nosologie zusammen. DSM-5 und ICD-10 bedienen sich fast ausschließlich syndromaler Kriterien zur Klassifikation psychischer Störungen. Die individuellen Verlaufsformen und das sehr variable Ansprechen auf Therapien legen nahe, dass sich hinter den heutigen diagnostischen Entitäten sehr heterogene pathophysiologische Prozesse verbergen (Casey et al., 2013; Kapur, Phillips & Insel, 2012; Krystal & State, 2014; Owen, 2014; Stephan et al., 2016a). Es fehlen der Psychiatrie bislang objektive Tests, welche pathophysiologische Prozesse im Individuum mess- oder sichtbar machen können, um gezieltere Aussagen zur Prognose oder Therapie zu machen.

Ein weiteres Problem ist, dass für verschiedene Störungsbilder zwar Erklärungsmodelle auf unterschiedlichen Ebenen (genetisch, molekular, kognitiv, etc.) bestehen, aber zumeist ein Modell fehlt, dass die verschiedenen Ebenen verbindet und mechanistische Erklärungen liefert, wie z. B. aus molekularen Veränderungen die klinisch

## Zentrale Themen der Computational Psychiatry

Die Nutzbarmachung mathematischer Methoden für die Lösung der obengenannten klinischen Probleme in der Psychiatrie steht im Zentrum der CP. Viele Befunde der neuropsychiatrischen Forschung beziehen sich auf Mittelwertsunterschiede zweier Gruppen. Da sich klinische Entscheidungsfindung jedoch immer auf individuelle Patienten bezieht, ist es für die mathematischen Modelle der CP zwangsläufig essentiell, auf Daten einzelner Patienten anwendbar zu sein und individuelle Krankheitsmechanismen erschliessen zu können. Diese Notwendigkeit schließt den Einsatz rein theoretischer Simulationsansätze aus und führt uns zu einem Kernkonzept der CP, die Entwicklung sogenannter „generativer Modelle“ für Daten, die im klinischen Kontext erhoben werden können, z. B. beobachtetes Verhalten oder mittels funktioneller Magnetresonanztomographie (fMRT) oder Elektroenzephalographie (EEG) gemessene Hirnaktivität.

### Generative Modelle als „Computational Assays“



**Abbildung 2.** Ein generatives Modell beschreibt, wie sich ein verborgener Systemzustand  $x$  auf einer nicht direkt beobachtbaren Systemebene in messbaren Daten  $y$  widerspiegelt, zum Beispiel die Umsetzung neuronaler Aktivität in ein messbares EEG-Signal. Diese Beschreibung beruht auf einem probabilistischen Vorwärtsmodell (likelihood: die Wahrscheinlichkeit  $p(y|x, m)$ , dass Messung  $y$  resultiert, wenn Zustand  $x$  vorliegt) und *a priori* Wahrscheinlichkeiten der Systemzustände  $p(x|m)$  (priors). Durch Anwendung des Bayes'schen Theorems lässt sich das Modell invertieren und die *a posteriori* Wahrscheinlichkeit  $p(y|x, m)$  des Systemzustands bei Vorliegen der gemessenen Daten berechnen. Diese Abbildung entstand aus zuvor publiziertem Bildmaterial von Garrido et al. (2008) und Stephan et al. (2016b), mit Genehmigung der Copyright-Besitzer.

## Generative Modelle und Computational Assays

Wie in Abbildung 2 dargestellt, beschreibt ein generatives Modell, wie sich ein latenter (d.h. verborgener) Prozess auf einer nicht direkt beobachtbaren Systemebene in messbaren Daten widerspiegelt. Zum Beispiel könnte ein generatives Modell im Bereich der Neurophysiologie beschreiben, wie sich die neuronale Aktivität in einer Region oder einem Schaltkreis in ein gemessenes EEG- oder fMRT-Signal umsetzt; ein anderes Beispiel aus dem behavioralen Bereich wäre ein generatives Modell, das abbildet, wie ein bestimmter kognitiver Algorithmus zu sichtbarem Verhalten führt.

Da Modelle per Definition eine abstrakte Vereinfachung darstellen und darüber hinaus Messungen stets mit Rauschen (noise) behaftet sind, ist ein generatives Modell stets probabilistisch: es sagt vorher, mit welcher Wahrscheinlichkeit bestimmte messbare Daten resultieren, wenn ein bestimmter Systemzustand vorliegt. Aus klinischer Anwendungssicht möchte man aber den gegenteiligen Schluss ziehen, d.h., man möchte wissen, welcher Systemzustand (z.B. ein verborgener Krankheitsprozess) vorliegt, wenn bestimmte Messdaten erhoben wurden. Dies lässt sich prinzipiell über die Anwendung eines zentralen Satzes der Wahrscheinlichkeitstheorie – das sogenannte Bayes'sche Theorem – erreichen. Durch die Umkehrung der Richtung der Modell-Vorhersage (Invertierung des Modells), lässt sich der verborgenen Systemzustand aus den Messdaten inferieren (probabilistisch erschliessen); dies wird auch als „Lösung des inversen Problems“ bezeichnet. Anders gesagt, man kann mit Hilfe Bayesianischer Methoden berechnen, mit welcher Wahrscheinlichkeit bei gegebenen Messdaten ein bestimmter Systemzustand (oder Krankheitsprozess) vorliegt. Entscheidend hierbei ist, dass ein generatives Modell *a priori* Wahrscheinlichkeiten (Priors) über die Parameter des verborgenen Systemzustands – z.B. die Stärke synaptischer Verbindungen – einbezieht.

Mithilfe eines geeigneten generativen Modells lassen sich somit prinzipiell aus nicht-invasiven Messungen bildgebender oder elektrophysiologischer Verfahren, aber auch aus einfachen Verhaltensmessungen (z.B. im Rahmen computerbasierter Aufgaben oder Spiele), Aussagen über krankheitsrelevante neurophysiologische oder kognitive Prozesse machen. So ist es zum Beispiel möglich, aus elektrophysiologischen Daten auf pharmakologisch oder genetisch induzierte Veränderungen synaptischer Übertragung zu schließen. Empirische Beispiele finden sich in Tier- und Humanstudien, in denen generative Modelle genutzt wurden, um Veränderungen in der Balance glutamaterger und GABAerger Aktivität unter variierender Anästhetiefe zu bestimmen (Moran et al., 2011a), den

Einfluss dopaminerger Medikation auf ionotrope (glutamaterge) Rezeptoren im Kortex zu ermitteln (Moran, Symmonds, Stephan, Friston & Dolan, 2011b), oder monogenetische Ionenkanalveränderungen zu detektieren (Gilbert et al., 2016).

Dieser Ansatz, mittels eines generativen Modells aus peripheren Messungen auf pathophysiologische und psychopathologische Prozesse zu schließen, wird auch als „Computational Assay“ bezeichnet (Stephan, Iglesias, Heinzle & Diaconescu, 2015; Stephan & Mathys, 2014) – in Analogie zu biochemischen Assays, mit denen aus peripheren Gewebeproben auf Krankheitsprozesse in inneren Organen geschlossen wird.

Im Folgenden werden wir in aller Kürze einige der wesentlichen generativen Modellierungsansätze vorstellen, die in der CP genutzt werden. Wir unterscheiden hierbei generative Modelle von Verhaltensdaten („komputationale Modelle“ im engeren Sinne, da sie die algorithmische Struktur von Informationsverarbeitung beschreiben) und generative Modelle von neurophysiologischen Daten.

## Generative Modelle von Verhalten

Stark vereinfacht lassen sich mathematische Modelle von Verhalten in zwei Klassen aufteilen, die sich vorrangig auf Entscheidungsfindung bzw. Wahrnehmung konzentrieren. Eine erste Klasse von Modellen, mit partiell bis zum Behaviorismus zurückreichenden Wurzeln, beruhen auf sogenannten Markov Decision Processes (MDP) (Sutton & Barto, 1998). Dieses Konzept sieht vor, dass Handlungen stets so gewählt werden, dass über einen bestimmten Zeitraum die größte Belohnung (reward) bzw. Nützlichkeit (utility) resultiert. Um dies zu erreichen, gilt es, in jedem Entscheidungskontext eine endliche Anzahl von relevanten Zuständen der Welt zu unterscheiden und den Wert (value) eines jeden Zustands bzw. den Wert der entsprechenden Handlungsoptionen zu lernen. Da es mit zunehmender Komplexität des Problems schnell unmöglich wird, die dem beobachteten Verhalten zu Grunde liegenden Entscheidungsprozesse mittels eines MDP exakt zu bestimmen, existieren praktikable Näherungsverfahren, wie z.B. das Verstärkungslernen (reinforcement learning, RL). Bestimmte RL-Modelle haben eine empirisch gezeigte Verbindung zur Physiologie, als dass die Aktivität dopaminerger Neurone im Mittelhirn (unter anderem) die Diskrepanz zwischen einem erwarteten und einem tatsächlichen Stimulus widerspiegelt (der sogenannte „Vorhersagefehler“ oder „prediction error“; Schultz, Gauthier & Montague, 1997).

RL-Modelle sind bereits bei etlichen psychiatrischen Erkrankungen eingesetzt worden, um mögliche Mechanismen des veränderten Verhaltens formal zu beschreiben

(Maia & Frank, 2011). So ergeben sich aus der Anwendung von RL-Modellen auf Verhaltensdaten von Patienten Hinweise darauf, dass der Anhedonie in der Depression bzw. Negativsymptomen in der Schizophrenie eine verminderte Empfindlichkeit für Belohnungserlebnisse zu Grunde liegen könnte (Huys, Pizzagalli, Bogdan & Dayan, 2013; Kirschner et al., 2016; Radua et al., 2015). Ein weiterer Anwendungsschwerpunkt liegt im Bereich der Abhängigkeitserkrankungen, wo Störungen dopaminerger Mechanismen bei verstärkendem Lernen vermutet werden (Deserno et al., 2015; Redish, Jensen, Johnson & Kurth-Nelson, 2007).

Eine zweite Klasse von Modellen – sogenannte hierarchische Bayesianische Modelle (HBM) wie Predictive Coding (Friston, 2005; Rao & Ballard, 1999) oder Hierarchical Filtering (Iglesias et al., 2013; Mathys, Daunizeau, Friston & Stephan, 2011) – beschäftigt sich mit dem fundamentalen Problem der Wahrnehmung (Perzeption): Das Gehirn hat keinen direkten Zugriff auf den Zustand der Welt – z.B. die physikalische Umwelt, die mentale Zustände anderer Menschen, der Zustand des eigenen Körpers – sondern kann diesen nur aus seinen Sinneseingängen probabilistisch, d.h. auf Basis von Wahrscheinlichkeiten, erschließen. Eine ideale Weise, dieses Inferenz-Problem zu lösen, leitet sich aus einem zentralen Satz der Wahrscheinlichkeitstheorie ab, dem oben erwähnten Satz von Bayes. Dieses Theorem liegt auch dem oben besprochenen Konzept des generativen Modells zu Grunde, beschreibt hier aber, wie sich im Sinne einer kognitiven Verarbeitung gegebenes Vorwissen oder Vorab-Erwartung (prior beliefs) mit neuen Sinneseindrücken kombinieren lässt, um optimale Schlüsse (posterior beliefs) zu ziehen.

Die Übertragung dieses optimalen Inferenzprinzips in den Bereich der Perzeption führte zu der inzwischen weitverbreiten „Bayesian Brain“ Theorie. Diese besagt, dass das Hirn ein Modell seiner Umwelt erstellt, welches vorhersagt, wie ein jeder gegebener Weltzustand zu sensorischen Inputs führt (Friston et al., 2014); also ein generatives Modell der sensorischen Eingänge. Durch Kombination von (aus gelernten und im Modell kodierten) Vorerwartungen und den sensorischen Daten kann das Hirn auf den verborgenen Zustand der Welt schließen (Inversion des generativen Modells). Diese Theorie steht nicht nur im Einklang mit einer Vielzahl an Befunden aus der Wahrnehmungspsychologie (Knill & Richards, 1996; Yuille & Kersten, 2006), sondern kann auch alltägliche (und nur scheinbar fehlerhafte) Aspekte menschlicher Perzeption wie Wahrnehmungsverzerrung oder sensorische Illusionen umfassend erklären (Petzschner, Glasauer & Stephan, 2015).

HBM bieten einen sehr allgemeinen Rahmen für die konkrete Implementierung dieser Theorie an, indem sie für das vom Gehirn erstellte generative Modell eine hierarchische Struktur annehmen (Friston, 2005; Mathys et al.,

2011; Rao & Ballard, 1999). Dies erlaubt, auch komplizierte Strukturen der Welt abzubilden (z.B. multiple zeitliche und räumliche Skalen, oder semantische Zusammenhänge) und steht im Einklang mit der bemerkenswert hierarchischen Struktur des Kortex (Felleman & Van Essen, 1991; Hilgetag, O'Neill & Young, 2000). Hierbei stellt eine jede Ebene des HBM eine probabilistische Repräsentation (belief) des erwarteten Zustands einer Variablen in der Welt dar. Durch den Austausch von Prädiktionssignalen und Vorhersagefehlern zwischen den Ebenen wird ein optimales Erschließen des Zustands der Welt ermöglicht. Hervorzuheben ist, dass HBM dabei auch abbilden, mit welcher subjektiven (Un)Gewissheit Zustände der Welt erschlossen werden.

HBM werden zurzeit von vielen Arbeitsgruppen zur Untersuchung und Formalisierung verschiedener psychiatrischer Symptome, aber auch zur Theoriebildung, eingesetzt. Drei Beispiele sollen hier in aller Kürze erwähnt werden. Im Bereich der Schizophrenie dienen HBM zur Erklärung von Halluzinationen (Fletcher & Frith, 2009; Friston, 2005; Horga, Schatz, Abi-Dargham & Peterson, 2014) und Wahnvorstellungen (Adams, Stephan, Brown, Frith & Friston, 2013; Corlett, Taylor, Wang, Fletcher & Krystal, 2010). In Bezug auf Angst bzw. Stressreaktionen konnten HBM in empirischen Studien (an gesunden Probanden) zeigen, dass emotionale und physiologische Stress-Antworten eng mit der subjektiven Ungewissheit bei der Vorhersage aversiver Stimuli korrelieren (de Berker et al., 2016) und dass Probanden mit stärker ausgeprägten ängstlichen Persönlichkeitszügen Schwierigkeiten haben, ihr Lernen an Ungewissheit über wechselnde Wahrscheinlichkeiten (Volatilität) anzupassen (Browning, Behrens, Jocham, O'Reilly & Bishop, 2015). Schließlich dienen HBM als Grundlage einer weitfassenden Theorie menschlicher Kognition und Entscheidungsfindung, die von Bedeutung für psychiatrische Krankheitskonzepte ist. Das sogenannte „Prinzip der freien Energie“ (free energy principle; Friston, 2010) besagt, stark vereinfacht gesagt, dass das Gehirn einem einzigen Ziel folgt: die Überraschung (im informationstheoretischen Sinne) über seine sensorischen Eingänge zu minimieren. Dies kann es auf zwei Arten erreichen: durch Anpassung seines generativen Modells (so dass die vorhergesagten Empfindungen besser zu den tatsächlichen passen), oder durch Auswahl von Handlungen, welche die erwarteten sensorischen Eingänge wahr werden lassen (active inference; Friston et al., 2015).

Diese allgemeine Idee ermöglicht die Konstruktion konkreter pathogenetischen Theorien mit möglichen Implikationen für Differenzialdiagnostik. Ein aktuelles Beispiel liefert eine auf HBM basierende metakognitive Theorie der Depression, die Patienten-Subgruppen mit differenziellen Störungen aktiver Inferenz vorhersagt (Stephan et al., 2016b). Gemäss dieser Theorie führen bei ei-



nem Teil depressiver Patienten Störungen der Interozeption bzw. allostatischer Kontrolle zu mangelhafter Erfüllung von *a priori* Erwartungen über Körperzustände (homeostatic beliefs). Die resultierende chronische Dyshomöostase wird auf einer metakognitiven Ebene als mangelnde Selbstwirksamkeit auf der fundamentalen Ebene der Selbsterhaltung erlebt, mit anschließender Generalisierung in depressive Erschöpfung und Antriebsmangel. Wichtig ist, dass diese Theorie durch die Anwendung von HBM auf neurophysiologische Daten von Patienten überprüfbar ist und, falls sie sich als stichhaltig erweisen sollte, direkt zu einem differentialdiagnostischen Hilfsinstrument auf der Basis eines Computational Assays führen könnte. Schliesslich würden sich durch die differentialdiagnostische Zuordnung zu subgruppenspezifischen Krankheitsmechanismen gezielte Behandlungsansätze ergeben, z. B. in der Psychotherapie (siehe unten).

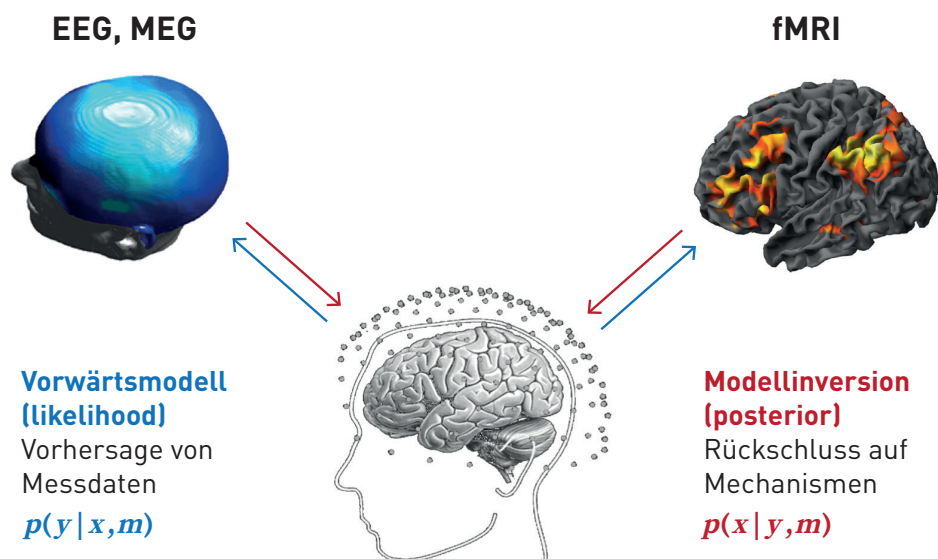
## Generative Modelle neurophysiologischer Daten

Generative Modelle für die Analyse von fMRT- und EEG-Daten lassen sich in zwei wesentliche Modellklassen unterscheiden: biophysikalische Netzwerkmodelle und Dynamic Causal Models (für eine aktuelle Übersichtsarbeit, siehe Stephan et al., 2015).

Biophysikalische Netzwerkmodelle (BNM) beruhen auf einer anatomisch und biophysikalisch detaillierten Beschreibung von großen (oft das ganze Gehirn abdeckenden) Netzwerken (Deco, Jirsa & McIntosh, 2013a). Die

einzelnen Knoten dieser Netzwerke repräsentieren lokale Areale oder Nuclei; die darin enthaltenen Neuronen werden nicht einzeln mathematisch abgebildet, stattdessen wird ihre Populationsaktivität modelliert. Die Verbindungen zwischen den Arealen sind durch anatomische Befunde aus tierexperimentellen Studien oder aus diffusionsgewichteten bildgebenden Studien im Menschen geleitet. Das Ausmaß an biophysikalischen und anatomischen Details, das BNM zu Grunde liegt, ist gleichermaßen Segen wie Fluch: einerseits ist die direkte Interpretierbarkeit dieser Modellklasse attraktiv, andererseits ist es numerisch bislang schwierig bis unmöglich, die sehr große Anzahl an freien Parametern in diesen Modellen aus Daten zu schätzen. Bislang können nur sehr wenige Parameter dieser Modelle bestimmt werden; zumeist handelt es sich hierbei um einen einzigen globalen Skalierungsparameter der Konnektivität (Deco et al., 2013b).

Eine zweite Modellklasse, die Dynamic Causal Models (DCMs; Friston, Harrison & Penny, 2003), basieren auf einer im Vergleich zu BNM vereinfachten Beschreibung der Aktivität interagierender neuronaler Populationen (Abb. 3). In Kombination mit fortgeschrittenen Bayesianischen Methoden ermöglicht es diese Vereinfachung, aus fMRT- und EEG-Daten einzelner Probanden Parameter zu schätzen, die individuelle Netzwerk-Mechanismen in der Form effektiver Konnektivität beschreiben. Um mathematische Handhabbarkeit zu gewährleisten, sind diese Modelle üblicherweise auf relativ kleine Netzwerke (ca. 10 Regionen) beschränkt. Die Erweiterung dieses Ansatzes auf Ganzhirnmodelle ist ein heißes Thema der aktuellen Forschung.



**Abbildung 3.** Schematische Darstellung von Dynamic Causal Modeling (DCM), einem generativen Modellierungsprinzip für fMRT- und EEG-Daten. Hier bezeichnet  $y$  gemessene Daten;  $x$  neuronale Zustände; und  $m$  das Modell. Vergleiche Abbildung 2. Diese Abbildung entstand aus zuvor publiziertem Bildmaterial von Chen, Henson, Stephan, Kilner & Friston (2009), Garrido et al. (2008) und Stephan et al. (2016c), mit Genehmigung der Copyright-Besitzer.

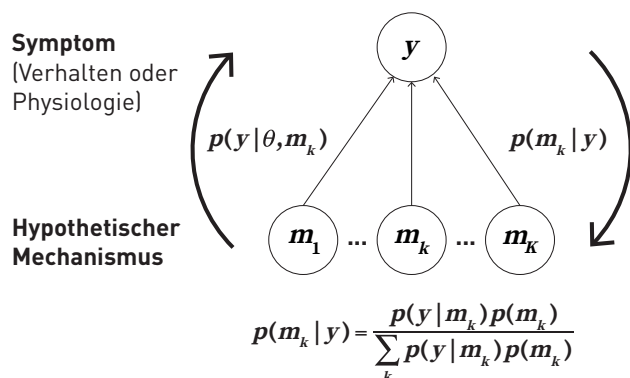
Hervorzuheben ist, dass DCMs für EEG-Daten bereits jetzt mit ähnlicher biophysikalischer Präzision wie BNMs formuliert werden können, wenn man sich auf die Untersuchung lokaler Schaltkreise konzentriert. Tierexperimentelle Validierungsstudien und pharmakologische Humanstudien haben gezeigt, dass dieser Ansatz auch die Bestimmung detaillierter physiologischer Mechanismen, wie z.B. die Veränderung der Exzitation-Inhibition-Balance (Moran et al., 2011a) oder neuromodulatorische Effekte (Moran et al., 2013) und ihren Einfluss auf glutamaterge NMDA-Rezeptoren (Moran et al., 2011b), ermöglicht.

## Modellselektion für Differentialdiagnostik

Über Probanden bzw. Patienten hinweg können anatomische oder physiologische Determinanten eines neuronalen Schaltkreises erheblich variieren, als Ausdruck unterschiedlicher genetischer Faktoren oder unterschiedlicher Ausprägung synaptischer Plastizität durch Umweltfaktoren und Erfahrung. Des Weiteren können sich Patienten deutlich in der Ausführung kognitiver Prozesse und dem daraus resultierenden Verhalten unterscheiden, wenn verschiedene algorithmische Strategien gewählt werden oder unterschiedliche Erwartungen (beliefs) vorliegen.

Beide Formen von Varianz sind von größtem klinischen Interesse, da sie direkt mit einem der zentralen Probleme der Psychiatrie (aber auch anderer klinischer Disziplinen)

## Differentialdiagnose durch Modellselektion



**Abbildung 4.** Ein Ziel der CP ist die Ermöglichung formaler Differentialdiagnose durch Modellselektion. Dies erfordert, dass man für ein gegebenes Symptom  $y$  (ein Verhaltens- oder physiologischer Befund) generative Modelle  $m_1 \dots m_k$  entwirft, von denen ein jedes einen alternativen Krankheitsprozess darstellt, um das Symptom zu erklären. Durch Anwendung des Satz von Bayes lassen sich die a posteriori Modell-Wahrscheinlichkeit  $p(m_k | y)$  für jedes Modell  $k$  bestimmen und entscheiden, welches Modell und damit welcher Krankheitsprozess am plausibelsten ist. Diese Abbildung entstand aus zuvor publiziertem Bildmaterial von Stephan et al. (2016c), mit Genehmigung der Copyright-Besitzer.

verbunden sind: unsere momentan eingeschränkte Fähigkeit, für Patienten maßgeschneiderte Therapieempfehlungen auf der Basis einer differenzialdiagnostischen Einsicht in individuelle Krankheitsmechanismen zu ermöglichen.

Eine aktuelle, auf dieses Problem ausgerichtete Forschungsstrategie ist die Entwicklung einer Batterie von Computational Assays, die alternative Krankheitsmechanismen über verschiedene generative Modelle abbilden (Stephan et al., 2016c). Durch statistische Modellvergleiche könnte das im einzelnen Patienten plausibelste Modell für die Erklärung der gemessenen physiologischen bzw. behavioralen Daten gefunden werden, um einen Hinweis auf die Natur des individuellen Krankheitsmechanismus zu erhalten. Dies würde eine formalisierte Differentialdiagnostik ermöglichen, mit der die Wahrscheinlichkeit verschiedener Erklärungsansätze (posterior model probability) bestimmt werden kann (Abb. 4). Wir gehen darauf unten näher ein.

## Maschinelles Lernen und Generative Embedding

Bereits frühe diagnostische Überlegungen in der Psychiatrie hatten zum Ziel diagnostische Gruppen zu umschreiben, die den Verlauf oder das Therapieansprechen vorherzusagen (vgl. z.B. die Dichotomie zwischen gutartig und ungünstig verlaufenden Erkrankungen in der Kraepelinschen Nosologie (Kraepelin, 1896). Neben der Suche nach einer auf Krankheitsmechanismen beruhenden Nosologie hat auch die statistische Suche nach prädiktiven Mustern in klinischen oder biologischen Daten eine lange Tradition in der psychiatrischen Forschung; für Beispiele aus den letzten Jahrzehnten, siehe Huxley, Goldberg, Maguire & Kincey (1979), Itil, Marasa, Saletu, Davis & Mucciardi (1975) und Nedopil, Pflieger & Ruther (1983). In jüngerer Zeit hat diese Arbeit durch Fortschritte im maschinellen Lernen (ML), das sich mit der statistischen Erkennung von Mustern in komplexen Datensätzen beschäftigt, Auftrieb erhalten (Klöppel et al., 2012). So konnte z.B. durch Anwendung moderner ML-Methoden gezeigt werden, dass sich aus rein klinischen Daten mit 59 % Genauigkeit das Ansprechen auf Antidepressiva vorhersagen lässt (Cherkoud et al., 2016), und aus strukturellen MRT-Bildern das individuelle Ansprechen depressiver Patienten auf Elektrokrampftherapie (Redlich et al., 2016) bzw. das Risiko von hochgefährdeten Individuen, den Übergang in die Psychose zu erleben mit je ca. 80 % Genauigkeit (Koutsouleris et al., 2015). Diese mit klassischen ML-Methoden erzielten Ergebnisse stellen wichtige Fortschritte auf dem Weg zu klinischen relevanten Werkzeugen darstellen dar, sind aber aus methodischer Sicht „black box“ Resultate, deren Prädiktion nicht direkt auf konkrete oder intuitiv

nachvollziehbare biologische oder kognitive Mechanismen zurückgeführt werden kann. Klassische ML-Methoden haben den Vorteil, dass sie nicht zwingend ein Verständnis der Daten erfordern und somit auch von fachfremden Wissenschaftlern angewendet werden können, haben aber den wichtigen Nachteil, keine mechanistische Interpretation zuzulassen und somit keine Orientierung für neue Therapieansätze zu liefern.

Ein Ansatz, die Vorteile maschinellen Lernens und generativer Modellierung zu verbinden, ist das sogenannte Generative Embedding (Brodersen et al., 2014; Brodersen et al., 2011). Bei diesem Ansatz wird die Fähigkeit eines generativen Modells, komplexe Daten in (wenige) Parameter und statistisches Rauschen zu separieren, genutzt, um dem anschließenden ML eine kompakte und mechanistische interpretierbare Zusammenfassung der Daten zu übergeben. Sofern ein plausibles generatives Modell existiert, kann dieser Ansatz nicht nur die Prädiktionsgenauigkeit und die Validität detektierter Muster (z. B. Cluster von Patienten) erhöhen, sondern erlaubt auch eine direkte Interpretierbarkeit der Resultate, da sich zurückverfolgen lässt, welcher der Parameterwerte (und damit welcher biologischer oder kognitiver Mechanismus) die erzielte Prädiktion möglich macht.

## Erhoffte Beiträge der CP zur zukünftigen klinischen Praxis in der Psychiatrie

In den vorangegangenen Abschnitten haben wir bereits etliche Beispiele erwähnt, wie die CP zukünftig zur klinischen Praxis in der Psychiatrie beitragen könnte. In diesem Abschnitt versuchen wir, die wesentlichen klinischen Ziele der CP kurz zusammenzufassen und zu systematisieren. Diese lassen sich in zwei Hauptgruppen unterteilen: die Entwicklung von Computational Assays als Hilfswerkzeuge zur Lösung konkreter klinischer Probleme, und die Ermöglichung konzeptueller Fortschritte in Nosologie und psychiatrischer Forschung. Es sollte dabei betont werden, dass die CP noch in den Kinderschuhen steckt und, wie weiter unten diskutiert, die Machbarkeit der erhofften klinischen Beiträge sich noch in prospektiven Validierungsstudien erweisen muss. Die folgenden Ausführungen sind also lediglich als ein hoffnungsvoller Entwurf für die Etablierung der CP zu verstehen.

Eine wichtige Strategie für CP, die ihre Inspiration aus Erfolgsgeschichten anderer klinischer Disziplinen zieht, ist die Entwicklung von Computational Assays, mit denen aus im klinischen Kontext gewinnbaren Daten auf Krankheitsprozesse im individuellen Patienten geschlossen werden

kann. Sollten sich Assays mit hinreichender Sensitivität und Spezifität als auch klinisch sinnvoller Interpretierbarkeit etablieren lassen, so könnte dies einen bedeutsamen Beitrag zur Lösung aktueller Probleme der psychiatrischen Praxis leisten. Darunter fallen:

### Verbesserte und beschleunigte Diagnose

Die Diagnosestellung gemäß der syndromalen Klassifikationssysteme wie DSM oder ICD ist zumeist nicht das dringendste klinische Problem. Allerdings gibt es bedeutende Ausnahmen, bei denen das Vorhandensein eines beobachterunabhängigen diagnostischen Tests eine erhebliche Erleichterung und Beschleunigung der Diagnostik bedeuten würde. Darunter fällt zum Beispiel die Diagnose der Autismusspektrumsstörung (ASS; siehe Haker, Schneebeli & Stephan, 2016) für eine eingehende Diskussion). Im Kindesalter erfordert diese Diagnose aufwändige diagnostische Abklärungen mittels semi-strukturierter Interviews und Beobachtungen (wie z. B. Autism Diagnostic Observation Schedule, ADOS, und Autism Diagnostic Interview, ADI-R); im Erwachsenenalter fehlen standardisierte Instrumente, und psychiatrische Expertise ist dünn gesät. Die beschleunigte Diagnosestellung einer ASS kann im Kindesalter Verzögerungen bei der Einleitung verhaltenstherapeutischer Maßnahmen vermeiden; darüber hinaus bedeutet die Diagnosestellung häufig eine große psychische Erleichterung für die Betroffenen und die Familien, die eine Erklärung für die beobachteten Symptome suchen. Ein anderes klinisches Beispiel betrifft die Unterscheidung von unipolaren und bipolaren affektiven Störungen, insbesondere bei Patienten, die sich initial mit depressiven Symptomen präsentieren. Hier könnte ein Test zur präzisen Diagnosestellung eine frühzeitige Behandlung mit stimmungsstabilisierender Medikation ermöglichen, um Manien und ihre sozialen Folgen zu verhindern.

### Differenzialdiagnose und prädiktive Tests auf individuelles Therapieansprechen

Wie oben beschrieben, könnte eine formale Differenzialdiagnose zwischen verschiedenen Typen von Krankheitsmechanismen durch den statistischen Vergleich verschiedener generativer Modelle ermöglicht werden. Alternativ könnte eine Differenzialdiagnose auf den geschätzten Parameterwerten eines Modells aufsetzen. In beiden Fällen ist Differenzialdiagnose nicht in Bezug auf DSM/ICD-Diagnosen zu verstehen, sondern in Bezug auf pathophysiologische oder psychopathologische Prozesse. In Bezug auf Pathophysiologie existieren zum Beispiel mehrere po-



tentielle EEG-basierte Computational Assays, die in initialen Validierungsstudien eine vielversprechende Fähigkeit zeigen, Veränderungen synaptischer Übertragung in spezifischen Netzwerken zu detektieren, z.B. glutamaterge und cholinerge Transmission in auditorischen-präfrontalen Schaltkreisen (Moran et al., 2013; Schmidt et al., 2013b), oder glutamaterge und dopaminerge Transmission im präfrontalen Kortex (Moran et al., 2011b).

Angesichts der Tatsache, dass die Mehrzahl der in der Psychiatrie genutzten Pharmaka die Synthese, Metabolismus oder postsynaptische Wirkung neuromodulatorischer Transmitter (wie Dopamin, Acetylcholin oder Serotonin) beeinflussen, könnten modell-basierte Indizes von synaptischer Übertragung und neuromodulatorischen Prozessen für die Planung individueller Pharmakotherapie nützlich sein. Diese Hypothese wird zurzeit in einer prospektiven Studie mit an Schizophrenie leidenden Patienten untersucht. Die in Zürich durchgeführte Beobachtungsstudie COMPASS ([www.compass-studie.ch](http://www.compass-studie.ch)) versucht vorherzusagen, ob Patienten, bei denen eine initiale antipsychotische Medikation ohne cholinerge Wirkkomponente keine ausreichende Wirkung zeigt, von einer Ergänzung bzw. Umstellung auf eine antipsychotische Medikation mit cholinerg (muskarinerger) Wirkkomponente, d.h. Clozapin oder Olanzapin, profitieren. Zu diesem Zwecke werden EEG- und fMRT-Messungen vor Beginn der Medikationsanpassung durchgeführt und eine klinische Nachbeobachtung über acht Wochen durchgeführt. Das Ziel ist zu testen, ob durch Computational Assays gewonnene Indizes synaptischer Übertragung in verschiedenen neuronalen Schaltkreisen das individuelle Ansprechen auf die Medikationsanpassung vorhersagen können.

Computational Assays könnten aber auch für psychopathologische Differenzialdiagnose nützlich sein. So implizieren die oben diskutierten HBM, dass Wahrnehmungsverzerrungen, die z.B. psychotischen Symptomen zu Grunde liegen mögen, durch Störungen verschiedener komputationaler Prozesse hervorgerufen werden könnten, d.h. der Berechnung oder Signalisierung von Vorabinformationen bzw. Erwartungen, Vorhersagefehlern oder ihrer Gewissheit bzw. Präzision (Stephan et al., 2016a). Patienten mit ähnlichen psychotischen Symptomen könnten diesbezüglich variieren, und eine differenzialdiagnostische Einordnung hätte mögliche Implikationen für einen verhaltenstherapeutischen Zugang. Ähnlich dem metakognitiven Training könnte durch Einsicht in die Entstehung verzerrter Wahrnehmung eine entsprechende Umstrukturierung angestrebt werden. Ein anderes Beispiel betrifft die Unterscheidung bei depressiven Patienten, ob eine Störung der Interozeption (mit subjektiv erlebter homöostatischer Dysregulation) vorliegt oder nicht. Hier könnte eine modellbasierte Diagnostik beitragen, neue Differentialtherapien, wie z.B. ein Körperwahrnehmungstraining

im Rahmen der psychotherapeutischen Interventionen für die Gruppe mit einer Störung der Interozeption, zu überprüfen.

## Konzeptuelle Fortschritte durch CP

CP hat nicht nur das Potential, zur Lösung konkreter klinischer Entscheidungsprobleme beizutragen, sondern auch konzeptuelle Fortschritte zu ermöglichen. Aus der Sicht der Patienten bietet sie Möglichkeiten zur Psychoedukation, indem sie neuartige Krankheitskonzepte anbietet, die eine subjektive Konzeptbildung und ein mögliches Verständnis der eigenen Erkrankung unterstützen können. Ein konkretes Beispiel liefert die „Bayesian Brain“ Krankheitstheorie der ASS, nach der autistische Symptome entgegen weit verbreiteter Meinung nicht primär im Bereich sozialer Kognition angesiedelt sind, sondern durch suboptimale Inferenz (Übergewichtung der Sensorik im Verhältnis zu unpräzisen Vorerwartungen oder Priors) im Bereich basaler Perzeption entstehen und erst sekundär soziale Kognition beeinflussen (Haker et al., 2016). Die klinische Erfahrung der Letztautorin dieses Artikels ist, dass diese Krankheitstheorie den Betroffenen ein umfassenderes Konzept und Verständnis für ihre Störung als andere Erklärungsansätze vermittelt. Auch im Rahmen psychotischer Störungen ist ein solcher Ansatz vorstellbar und vielversprechend, als dass er sowohl psychoedukative sowie therapeutische Möglichkeiten impliziert. So kann in einem ersten Schritt durch die Verfügbarkeit eines Erklärungsansatzes für die Entstehung des Wahns eine subjektive Erleichterung eintreten. In einem weiteren Schritt könnte dann versucht werden, diese Erklärung gezielt einzusetzen, um eine kognitive Umstrukturierung zu erreichen.

Ein zweiter Bereich, in dem CP konzeptuelle Fortschritte unterstützen kann, betrifft die Brückenbildung über verschiedene Erklärungsebenen hinweg. Die meisten psychiatrischen Forschungskonzepte konzentrieren sich auf eine bestimmte Beschreibungsebene, z.B. genetische Ansätze auf die molekulare und psychopathologische Ansätze auf die kognitive Ebene. CP hat das Potenzial, diese Ebenen miteinander zu verbinden, in dem sie untersucht, wie neurophysiologische Mechanismen bestimmte komputationale Prozesse ermöglichen, die sich wiederum in Handeln und Wahrnehmung niederschlagen (Corlett & Fletcher, 2014).

Ein letzter, aber nicht minder wichtiger Punkt betrifft die Reform der psychiatrischen Nosologie, von einer syndromorientierten zu einer auf physiologischen und kognitiven Mechanismen beruhenden Klassifikation. Auch hier könnten generative Modelle einen wichtigen Beitrag leisten, indem sie helfen, heterogene Spektrumserkrankun-

gen wie Schizophrenie oder Depression in Subgruppen aufzulösen, die durch bestimmte physiologische oder kognitive Prozesse charakterisiert sind. Ein empirisches Beispiel liefert die Arbeit von Brodersen et al. (Brodersen et al., 2014), in der Generative Embedding genutzt wird, um eine Gruppe von 41 Patienten mit der Diagnose Schizophrenie in physiologisch definierte Untergruppen aufzuteilen, ohne dass Vorannahmen über die Struktur oder die Anzahl der Untergruppen getroffen wurden. Zu diesem Zweck wurde ein einfaches DCM auf fMRT-Daten einer Arbeitsgedächtnis-Aufgabe angewendet und die erhaltenen Konnektivitätsparameter zwischen kortikalen Regionen für eine Cluster-Analyse genutzt. Diese zeigte, dass die Verteilung der Konnektivitätsmuster am plausibelsten durch die Existenz dreier Untergruppen erklärt wurde. Bemerkenswerterweise unterschieden sich diese drei rein physiologisch definierten Subgruppen bezüglich der Ausprägung negativer Symptome, ohne dass das Modell Zugang zu klinischen Daten gehabt hätte. Dieses Beispiel zeigt, wie generative Modelle prinzipiell genutzt werden könnten, um die oft bedeutsame klinische Heterogenität von Patienten mit der gleichen DSM/ICD-Diagnose auf die unterschiedliche Ausprägung bestimmter biologischer oder kognitiver Mechanismen zu erklären. Langfristig könnte dies einen wichtigen Beitrag zu der seit langem angestrebten Reform psychiatrischer Krankheitsklassifikationen leisten (Insel, 2014; Owen, 2014; Stephan et al., 2016a).

## Grenzen und Herausforderungen für die Computational Psychiatry

Wie schon oben angedeutet, ist die CP eine extrem junge Disziplin, die viele methodische, aber auch organisatorische Herausforderungen erst noch bewältigen muss, bevor sie zu einer echten und verlässlichen Bereicherung der klinischen Praxis in der Psychiatrie führen kann. So ist ein wichtiges methodisches Problem, die existierenden Kandidaten-Assays in geeigneten Studien zu validieren und ihre Sensitivität und Spezifität für die Detektion bestimmter pathophysiologische Prozesse zu ermitteln. Diese Validierung ist bislang in nur sehr begrenztem Maße erfolgt. Generell benötigt die Validierung von modellbasierten klinischen Werkzeugen (i) Tierstudien mit möglichst selektiven Perturbationen krankheitsrelevanter Prozesse (für Beispiele im Kontext von DCM, siehe Moran et al., 2015; Moran et al., 2011a; Moran et al., 2008), (ii) pharmakologische Humanstudien (Moran et al., 2013; Moran et al., 2011b; Schmidt et al., 2013a), aber vor allem auch (iii) prospektive Patientenstudien, in denen geprüft wird,

ob die durch Computational Assays geschätzten Mechanismen überhaupt für individuelle Prognose bzw. Therapie-Prädiktion nützlich sind. Studien aller drei Formen sind schwierig durchzuführen, da sie langwierig, schwer planbar und nur im Rahmen interdisziplinärer Kooperationen möglich sind. Es wird wichtig sein, die wenigen vorhandenen Datensätze dieser Art systematisch für Re-Analysen (z.B. mit neu entwickelten Modellen) zur Verfügung zu stellen; dies wird ohne Anreizsysteme zum Datenteilen nur schwer möglich sein. Darüber hinaus gibt es nicht-triviale methodische Probleme bei der Perturbation (experimentellen „Störung“) physiologischer Prozesse in basalen Validierungsstudien in Tier und Mensch, z.B. die oft nicht im Detail bekannten Unterschiede zwischen akuten Perturbationen (im Experiment) und chronischen Veränderungen (bei einer Erkrankung) eines physiologischen Prozesses, oder die mangelhafte Selektivität der meisten uns zur Verfügung stehenden pharmakologischen Substanzen.

Insgesamt muss davon ausgegangen werden, dass es kaum realistisch sein wird, binnen weniger Jahre Computational Assays für die Psychiatrie zu entwickeln, welche die gleiche Auflösung und Trennschärfe haben, wie diagnostische Assays in der Inneren Medizin. Dies ist für diagnostische Nützlichkeit aber auch nicht zwingend nötig. Bereits eine Unterscheidung wesentlicher Typen von Mechanismen könnte, wie oben angedeutet, für die Therapieprädiktion von großer Nützlichkeit sein.

Nicht nur die Validierung, sondern auch die praktische Anwendung von Computational Assays steht erheblichen Herausforderungen gegenüber. Das Problem ist hierbei nicht in erster Linie die Verfügbarkeit von Software. Erste Open Source Programme beginnen von verschiedenen Instituten entwickelt zu werden, darunter das Zürcher Programmpaket TAPAS (<http://translationalneuromodeling.org/tapas>). Wie oben angedeutet, ist die Parameterschätzung in komplexen mathematischen Modellen kein triviales Unterfangen und erhaltene Resultate müssen sorgfältig auf ihre Plausibilität, Robustheit und potentielle numerische Probleme untersucht werden. Dies impliziert die Notwendigkeit einer kontinuierlichen Qualitätssicherung, die methodisch-orientierte Institute mit hochspezialisierter Expertise im Neuromodeling erfordert. Hier bietet sich die Analogie zur historischen Entwicklung der Inneren Medizin an, aus der sich Spezialgebiete wie die Klinische Chemie herauslösten und zu eigenen Disziplinen wurden, um die Qualität diagnostischer Werkzeuge zu gewährleisten.

Abschließend möchten wir auf ein besonders wichtiges Problem für die weitere Entwicklung der CP eingehen. Dies ist die hochgradig interdisziplinäre Natur dieser jungen Disziplin, die extrem unterschiedliche akademische Ansätze in sich vereint – mathematische Modellierungs-

techniken einerseits und biomedizinische Konzepte andererseits. Es wird für die Entwicklung des Felds unerlässlich sein, systematische Curricula zu definieren, in denen Lehrinhalte der Mathematik und Informatik gleichberechtigt klinischen Inhalten gegenüberstehen. Ein erster Ansatz dazu findet sich im Computational Psychiatry Course (CPC), ein seit 2015 jährlich von der Translational Neuromodeling Unit (TNU) Zürich durchgeführter Kurs, der regelmäßig mehr als 150 Studenten aus aller Welt anzieht und mathematische Modellierung mit klinischen Inhalten aber auch konkreter Einweisung in Software zu verbinden versucht (<http://www.translationalneuromodeling.org/cpcourse>).

Sollte es gelingen, die hier nur in aller Kürze dargestellten Probleme und Herausforderungen erfolgreich zu meistern, könnte sich die CP zu einem Teilbereich der klinischen Psychiatrie entwickeln, die in Zusammenarbeit mit theoretischen Neuromodeling-Instituten die zukünftige Praxis der Psychiatrie mit mathematischen Modellen für Differenzialdiagnose, Prognose und Therapie-Prädiktion bereichert.

## Danksagung

Wir danken der René und Susanne Braginsky Stiftung und der Universität Zürich für ihre Unterstützung.

## Literatur

- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D. & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, 4, 47.
- Brodersen, K. H., Deserno, L., Schlagenhaut, F., Lin, Z., Penny, W. D., Buhmann, J. M. & Stephan, K. E. (2014). Dissecting psychiatric spectrum disorders by generative embedding. *Neuroimage Clinical*, 4, 98–111.
- Brodersen, K. H., Schofield, T. M., Leff, A. P., Ong, C. S., Lomakina, E. I., Buhmann, J. M. & Stephan, K. E. (2011). Generative embedding for model-based classification of fMRI data. *PLoS Comput Biol*, 7, e1002079.
- Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, 18, 590–596.
- Casey, B. J., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S. & Ressler, K. J. (2013). DSM-5 and RDoC: progress in psychiatry research? *Nature Reviews Neuroscience*, 14, 810–814.
- Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., ... Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry*, 3, 243–250.
- Chen, C. C., Henson, R. N., Stephan, K. E., Kilner, J. M. & Friston, K. J. (2009). Forward and backward connections in the brain: a DCM study of functional asymmetries. *Neuroimage*, 45, 453–462.
- Corlett, P. R. & Fletcher, P. C. (2014). Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry*, 1, 399–402.
- Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C. & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in Neurobiology*, 92, 345–369.
- de Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J. & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, 7, 10996.
- Deco, G., Jirsa, V. K. & McIntosh, A. R. (2013a). Resting brains never rest: computational insights into potential cognitive architectures. *Trends in Neuroscience*, 36, 268–274.
- Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P. & Corbetta, M. (2013b). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, 33, 11239–11252.
- Deserno, L., Beck, A., Huys, Q. J., Lorenz, R. C., Buchert, R., Buchholz, H. G., ... Heinz, A. (2015). Chronic alcohol intake abolishes the relationship between dopamine synthesis capacity and learning signals in the ventral striatum. *European Journal of Neuroscience*, 41, 477–486.
- Felleman, D. J. & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1–47.
- Fletcher, P. C. & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10, 48–58.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society London*, 360, 815–836.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11, 127–138.
- Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T. & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, 6, 187–214.
- Friston, K. J., Harrison, L. & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, 19, 1273–1302.
- Friston, K. J., Stephan, K. E., Montague, R. & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry*, 1, 148–158.
- Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T. & Kilner, J. M. (2008). The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage*, 42, 936–944.
- Gilbert, J. R., Symmonds, M., Hanna, M. G., Dolan, R. J., Friston, K. J. & Moran, R. J. (2016). Profiling neuronal ion channelopathies with non-invasive brain imaging and dynamic causal models: Case studies of single gene mutations. *Neuroimage*, 124, 43–53.
- Haker, H., Schneebeli, M. & Stephan, K. E. (2016). Can Bayesian theories of autism spectrum disorder help improve clinical practice? *Frontiers in Psychiatry*, 7, 107.
- Hilgetag, C. C., O'Neill, M. A. & Young, M. P. (2000). Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor. *Philosophical Transactions of the Royal Society London*, 355, 71–89.
- Horga, G., Schatz, K. C., Abi-Dargham, A. & Peterson, B. S. (2014). Deficits in predictive coding underlie hallucinations in schizophrenia. *Journal of Neuroscience*, 34, 8072–8082.
- Huxley, P. J., Goldberg, D. P., Maguire, G. P. & Kinney, V. A. (1979). The prediction of the course of minor psychiatric disorders. *British Journal of Psychiatry*, 135, 535–543.
- Huys, Q. J., Maia, T. V. & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19, 404–413.
- Huys, Q. J., Pizzagalli, D. A., Bogdan, R. & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biological Mood and Anxiety Disorders*, 3, 12.



- Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., den Ouden, H. E. & Stephan, K. E. (2013). Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. *Neuron*, 80, 519–530.
- Insel, T. R. (2014). The NIMH research domain criteria (RDoC) project: precision medicine for psychiatry. *American Journal of Psychiatry*, 171, 395–397.
- Itil, T. M., Marasa, J., Saletu, B., Davis, S. & Mucciardi, A. N. (1975). Computerized EEG: predictor of outcome in schizophrenia. *Journal of Nervous and Mental Disorders*, 160, 118–203.
- Kapur, S., Phillips, A. G. & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17, 1174–1179.
- Kirschner, M., Hager, O. M., Bischof, M., Hartmann-Riemer, M. N., Kluge, A., Seifritz, E., ... Kaiser, S. (2016). Deficits in context-dependent adaptive coding of reward in schizophrenia. *NPJ Schizophrenia*, 2, 16020.
- Klöppel, S., Abdulkadir, A., Jack, C. R., Jr., Koutsouleris, N., Mourao-Miranda, J. & Vemuri, P. (2012). Diagnostic neuroimaging across diseases. *Neuroimage*, 61, 457–463.
- Knill, D. & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge: University Press.
- Koutsouleris, N., Riecher-Rössler, A., Meisenzahl, E. M., Smieskova, R., Studerus, E., Kambeitz-Ilankovic, L., ... Borgwardt, S. (2015). Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. *Schizophrenia Bulletin*, 41, 471–482.
- Kraepelin, E. (1896). *Psychiatrie* (5. Aufl.). Leipzig: J. A. Barth.
- Krystal, J. H. & State, M. W. (2014). Psychiatric disorders: diagnosis to therapy. *Cell*, 157, 201–214.
- Maia, T. V. & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, 14, 154–162.
- Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, 5, 39.
- Montague, P. R., Dolan, R. J., Friston, K. J. & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, 16, 72–80.
- Moran, R. J., Campo, P., Symmonds, M., Stephan, K. E., Dolan, R. J. & Friston, K. J. (2013). Free energy, precision and learning: the role of cholinergic neuromodulation. *Journal of Neuroscience*, 33, 8227–8236.
- Moran, R. J., Jones, M. W., Blockeel, A. J., Adams, R. A., Stephan, K. E. & Friston, K. J. (2015). Losing control under ketamine: suppressed cortico-hippocampal drive following acute ketamine in rats. *Neuropsychopharmacology*, 40, 268–277.
- Moran, R. J., Jung, F., Kumagai, T., Endepols, H., Graf, R., Dolan, R. J., ... Tittgemeyer, M. (2011a). Dynamic causal models and physiological inference: a validation study using isoflurane anaesthesia in rodents. *PLoS One*, 6, e22790.
- Moran, R. J., Stephan, K. E., Kiebel, S. J., Rombach, N., O'Connor, W. T., Murphy, K. J., ... Friston, K. J. (2008). Bayesian estimation of synaptic physiology from the spectral responses of neural masses. *Neuroimage*, 42, 272–284.
- Moran, R. J., Symmonds, M., Stephan, K. E., Friston, K. J. & Dolan, R. J. (2011b). An in vivo assay of synaptic function mediating human cognition. *Current Biology*, 21, 1320–1325.
- Nedopil, N., Pflieger, R. & Ruther, E. (1983). The prediction of acute response, remission and general outcome of neuroleptic treatment in acute schizophrenic patients. *Pharmacopsychiatry*, 16, 201–205.
- Owen, M. J. (2014). New Approaches to Psychiatric Diagnostic Classification. *Neuron*, 84, 564–571.
- Petzschner, F. H., Glasauer, S. & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Science*, 19, 285–293.
- Radua, J., Schmidt, A., Borgwardt, S., Heinz, A., Schlagenhauf, F., McGuire, P. & Fusar-Poli, P. (2015). Ventral striatal activation during reward processing in psychosis: A neurofunctional meta-analysis. *JAMA Psychiatry*, 72, 1243–1251.
- Rao, R. P. & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2, 79–87.
- Redish, A. D., Jensen, S., Johnson, A. & Kurth-Nelson, Z. (2007). Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychological Review*, 114, 784–805.
- Redlich, R., Opel, N., Grotegerd, D., Dohm, K., Zaremba, D., Burger, C., ... Dannlowski, U. (2016). Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry*, 73, 557–564.
- Schmidt, A., Diaconescu, A. O., Komter, M., Friston, K. J., Stephan, K. E. & Vollenweider, F. X. (2013a). Modeling ketamine effects on synaptic plasticity during the mismatch negativity. *Cerebral Cortex*, 23, 2394–2406.
- Schmidt, A., Smieskova, R., Aston, J., Simon, A., Allen, P., Fusar-Poli, P., ... Borgwardt, S. (2013b). Brain connectivity abnormalities predating the onset of psychosis: correlation with the effect of medication. *JAMA Psychiatry*, 70, 903–912.
- Schultz, W., Dayan, P. & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593–1599.
- Stephan, K. E., Bach, D. R., Fletcher, P. C., Flint, J., Frank, M. J., Friston, K. J., ... Breakspear, M. (2016a). Charting the landscape of priority problems in psychiatry, part 1: classification and diagnosis. *Lancet Psychiatry*, 3, 77–83.
- Stephan, K. E., Iglesias, S., Heinzle, J. & Diaconescu, A. O. (2015). Translational perspectives for computational neuroimaging. *Neuron*, 87, 716–732.
- Stephan, K. E., Manjaly, Z. M., Mathys, C., Weber, L. A. E., Paliwal, S., Gard, T., ... Petzschner, F. H. (2016b). Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience* (in revision).
- Stephan, K. E. & Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinions in Neurobiology*, 25, 85–92.
- Stephan, K. E., Schlagenhauf, F., Huys, Q. J., Raman, S., Aponte, E. A., Brodersen, K. H., ... Heinz, A. (2016c). Computational neuroimaging strategies for single patient predictions. *Neuroimage* (im Druck).
- Sutton, R. S. & Barto, A. G. (1998). Reinforcement learning: an introduction. *IEEE Trans Neural Network*, 9, 1054.
- Wang, X. J. & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, 84, 638–654.
- Yuille, A., Kersten, D., 2006. Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Science*, 10, 301–308.

#### Klaas Enno Stephan

Translational Neuromodeling Unit (TNU)  
Institut für Biomedizinische Technik  
Universität Zürich & ETH Zürich  
8032 Zürich  
Schweiz

stephan@biomed.ee.ethz.ch